

The role of unbalanced probabilities of occurrence in the identification of elementary auditory displays

I. D. John

University of Adelaide, Adelaide, South Australia 5001

Six groups of 20 Ss made absolute judgments of seven auditory stimuli differing in loudness. The experimental treatments for each group involved systematic manipulation of relative stimulus frequency (RSF), prior information about RSF, and knowledge of results. Pollack & Boynton's (1963) finding that identification performance was enhanced for the most frequently presented stimulus and depressed for adjacent stimuli was confirmed, but it was concluded that this finding resulted from a failure of Ss to match response frequency to the frequency of the most commonly presented stimulus even with knowledge of results and/or prior information about RSF. There was no evidence that increasing RSF resulted in more adequate learning of the most frequently presented stimulus.

It has been suggested by Eriksen & Hake (1957) that, in absolute judgment tasks, Ss use the end stimuli of the stimulus series as reference points for judging other stimuli in the series, and that this strategy results in a characteristic inverted U-shape relation between stimulus value and performance indicators such as number of errors and measures of response dispersion. Such a relationship has also been observed in the results of both category rating and magnitude estimation tasks (John, 1969), so there is some reason to believe that it may be common to all judgmental tasks involving a series of unidimensional stimuli.

An apparent exception to this general relationship has been reported in a study by Pollack & Boynton (1963). In their experiment, "listeners identified sound levels from a series in which one sound level occurred twice as often as each of six other sound levels. Unbiased measures of identification performance indicated that the highest identification performance is achieved for the most frequently presented category and that lowest identification performance is achieved for categories adjacent to the most frequent category." In their experiment Ss were provided with knowledge of results (KR) and prior information (PI) about relative stimulus frequency (RSF), and it is by no means clear whether the obtained results are due to the manipulation of RSF or to the influence of one or both of the other two factors.

The experiment to be reported was designed to elucidate Pollack and Boynton's findings. Three independent variables, RSF, KR, and PI were systematically manipulated to provide six experimental treatments, and measures of identification

performance were supplemented by measures of dispersion and relative frequency of responses.

METHOD

The stimulus on each trial consisted of a 2-sec presentation of one of seven intensities of a 1,000-Hz tone presented through a loudspeaker. The individual tones varied in intensity by steps of 5 dB over a range of 30 dB. The weakest intensity, which was designated "one," was clearly audible in all parts of the experimental room.

The Ss were drawn from an introductory psychology class, and the experimental procedure was administered in groups of 2-20 Ss; 20 Ss received each experimental treatment.

Details of the six experimental treatments are contained in Table 1. PI was provided (when appropriate) by instructing Ss about RSF prior to the commencement of the experiment. In the case of Group 6, false PI was given by stating that RSF for stimulus "four" would be .25 and for all other stimuli .125; in fact all stimuli had a RSF of .143. KR was provided (when appropriate) at the completion of each trial.

The manipulation of RSF was provided, in the case of four groups, by presenting the middle stimulus of the series (S4) on twice as many trials as each of the other six stimuli, to give RSFs of .25 and .125.

At the commencement of the experiment, all groups were provided with appropriate instructions and the stimulus series was demonstrated once in ascending and descending orders of presentation. The results of the first experimental trial were discarded from the analysis for all groups; the remaining 128 trials for groups with unequal RSF and 98 trials for groups with equal RSF were arranged as

follows. The trials were divided into two blocks of equal numbers of trials and the order of trials within blocks was randomized with the constraint that the relative frequency of sequences of two stimuli in two successive trials was equal to the product of the RSF of those two stimuli. The order of presentation of trials was therefore balanced with respect to the effects of the immediately preceding stimulus.

RESULTS

A summary of the relevant results is contained in Table 1. The data were first examined to determine if any constant error of judgment occurred. In the absence of such a constant error, the mean response should be 4.0 for each group. In Groups 2, 3, and 4, the mean response does not depart significantly from that value, but for Groups 1, 5, and 6, the mean response is significantly less than 4.0 ($p < .01$ in each case). A constant negative error therefore occurs in the absence of KR or accurate PI. To counterbalance the effects of the constant error, the remaining analyses were carried out in terms of performance associated with stimulus and response "four" (S4 and R4), and the mean performance associated with stimuli and responses "three" and "five," that is, the adjacent stimuli and responses (AS and AR).

Two accuracy measures, normalized proportions of stimulus presentations correctly identified and responses correctly assigned, were obtained. The normalization (Pollack & Boynton, 1963) is such that zero is tied to chance performance level based upon marginal sums and 1.00 is tied to completely correct performance. Two measures of dispersion were also obtained, response uncertainty conditional upon the individual stimulus [$U_s(R)$; Garner, 1962] and stimulus uncertainty conditional upon the individual response [$U_r(S)$]. Relative frequencies of response were also determined.

All differences between S4 and AS and R4 and AR for the measures of accuracy and dispersion were evaluated by two-tailed *t* tests with 19 df.

Group 1, for which RSFs are equal and for which KR and PI were not provided, may be regarded as a control group. In this group the accuracy measures are significantly less for S4 and R4 than for AS and AR, a finding which is consistent with the usual U-shaped relation between stimulus value and accuracy of performance. The differences between the dispersion measures are not significant for this group.

Group 2 is essentially a replication of Pollack and Boynton's experiment, and the results are consistent with

Table 1
Summary of Experimental Treatments and Performance Measures

Experimental Treatment			Results																
Group	KR	PI	Relative S Frequency		Relative R Frequency		\bar{X} Respon- ses	Proportion Correct S			Proportion Correct R			Us _i (R)			Ur _i (S)		
			S4	AS	R4	AR		S4	AS	SIG	R4	AR	SIG	S4	AS	SIG	R4	AR	SIG
1	No	No	.143	.143	.140	.154	3.61	.138	.240	**	.148	.206	*	1.45	1.46	n.s.	1.60	1.46	n.s.
2	Yes	Yes	.250	.125	.213	.134	3.94	.473	.375	*	.522	.360	***	1.54	1.58	n.s.	1.32	1.52	**
3	No	Yes	.250	.125	.216	.146	3.96	.426	.310	**	.442	.263	***	1.61	1.55	n.s.	1.49	1.65	**
4	Yes	No	.250	.125	.190	.139	3.99	.408	.364	n.s.	.541	.329	***	1.66	1.58	n.s.	1.29	1.57	**
5	No	No	.250	.125	.146	.155	3.52	.143	.229	**	.244	.177	n.s.	1.59	1.42	**	1.53	1.46	n.s.
6	No	False	.143	.143	.193	.135	3.69	.333	.202	*	.211	.207	n.s.	1.38	1.47	n.s.	1.65	1.52	*

*p < .05, **p < .01, ***p < .001, n.s. = not significant (two-tailed t test, df = 19).

their findings, namely that the accuracy measures are significantly greater for S4 and R4 than for AS and AR. This pattern of findings is repeated for both Groups 3 and 4 (although the difference between the two stimulus accuracy measures fails to reach significance for Group 4). This pattern is not, however, found in Group 5. It seems, therefore, that the enhancement of the accuracy measures for S4 and R4 and the reduction for AS and AR which was observed by Pollack and Boynton is not attributable to manipulation of RSF per se, but in addition requires either KR or accurate PI.

Some light is thrown on the operation of KR and PI in this context if the relative response frequencies are examined. In the absence of either KR or accurate PI (Groups 1 and 5), there seems to be a tendency for all responses to be used with the same relative frequency. In Groups 1 and 5 there is no significant difference between the relative frequencies of R4 and AR.

When RSF is manipulated, Ss are unable to change response frequencies appropriately; they seem to compromise between a frequency of responses matched to the frequency of stimuli and equal frequency of all responses. The significance of this trend has been assessed by expressing relative response frequencies for Groups 2, 3, and 4 as a proportion of RSF. In the case of all three groups, the proportions for AR are significantly greater than the proportions for R4 (p < .01). This tendency can also be observed in Group 6 when false PI has been provided; the relative response frequency for R4 in this group is .193 rather than .250, which would have occurred if frequency of responses was matched to PI about RSF.

The failure to match response frequency to stimulus frequency of the most frequently occurring stimulus may be thought of as a relative reduction of response strength. This failure of necessity produces an increase in frequency of ARs, which

may similarly be thought of as a relative increase in response strength. This interpretation is borne out by an examination of the Ur_i(S) measures. In the case of Group 1, where relative response frequencies of R4 and AR (and by implication, relative response strengths) are not significantly different, there is no significant difference between the Ur_i(S) measures. In Groups 2, 3, and 4, where the relative response strength of AR has been increased, Ur_i(S) associated with these responses is significantly greater than that associated with R4.

That is to say, increased response strength is characterized not only by an increased frequency of that response, but also greater inconsistency in the use of that response.

It is important to notice that the experimental treatment in Groups 2, 3, and 4 has not produced any relative changes in Us_i(R). Thus, there is no evidence that enhanced accuracy for S4 and R4 is in any way a reflection of learning of S4 or an increased ability to recognize S4.

The relative changes in accuracy following manipulation of RSF found by Pollack and Boynton thus appear to be attributable to changes in the response system, namely a failure to match frequency of responses to frequency of stimuli. A corollary of this matching failure is that those responses which are used more frequently than is justified by the number of corresponding stimuli are assigned more inconsistently than those responses which are used less frequently than is justified by the number of corresponding stimuli.

This explanation of Pollack and Boynton's findings is also consistent with a discrepancy which they pointed out between the results of their study and the previous study by Pollack (1963), in which the intelligibility of words within small known message sets was examined as a function of the probability of occurrence of the individual words. In this study identification performance was enhanced for the most frequent

stimulus but without a reduction in the performance for other stimuli. It seems clear that the essential difference between the two studies is that in the sound-identification task there is an ordered response continuum in which changes in relative response frequency for any response must be compensated for by changes in the relative response frequency of adjacent responses. Such compensatory changes in relative frequency of specific responses are by no means necessitated in the use of nonordered response continua.

It may be concluded, therefore, that the findings of this experiment are consistent with Eriksen and Hake's subjective standard hypothesis. They provide no evidence that stimuli within the stimulus series can be endowed with the same functional significance as the extreme stimuli of the series by means of a learning procedure such as has been employed in this experiment.

It should be noted, however, that more systematic and prolonged training procedures, such as those employed by Cuddy (1968) in an investigation of pitch discrimination, may give rise to genuine increases in the discriminability of an individual stimulus within a stimulus series so that it may be effectively utilized as a subjective standard.

REFERENCES

- CUDDY, L. L. Practice effects in the absolute judgment of pitch. *Journal of the Acoustical Society of America*, 1968, 43, 1069-1076.
- ERIKSEN, C. W., & HAKE, H. W. Anchor effects in absolute judgments. *Journal of Experimental Psychology*, 1957, 52, 132-138.
- GARNER, W. R. *Uncertainty and structure as psychological concepts*. New York: Wiley, 1962.
- JOHN, I. D. Stimulus discriminability in the magnitude estimation and category rating of loudness. *Perception & Psychophysics*, 1969, 2, 78-80.
- POLLACK, I. Message probability and message reception. *Journal of the Acoustical Society of America*, 1963, 35, 807-808.
- POLLACK, I., & BOYNTON, J. Identification of elementary auditory displays: Effects of unbalanced probabilities of occurrence. *Journal of the Acoustical Society of America*, 1963, 35, 1831-1832.